Realtime publishers

The Shortcut Guide™ To

Assuring Website Performance through External Web Monitoring

sponsored by

neustar™

Dan Sullivan

Copyright Statement

Chapter 3: Acquiring Customers, Controlling Costs, and Protecting Your Brand with External Monitoring

When you conduct business on the Web, your customer's browser becomes your store front. Businesses go to significant lengths to keep up appearances around their physical store fronts and offices. Buildings are cleaned and well decorated, parking lots are swept and well lit, and grounds are professionally landscaped. The outward appearance of a physical business reflects on the entire business itself. A broken door, a litter-riddled parking lot, and overgrown shrubs can make more of a negative impression than could be undone with desired characteristics, such as low-priced quality products. Customers, perhaps without being aware of it, expect a certain level of professionalism and quality in the delivery of services. This is true whether you are talking about physical stores or online services.

Poor performance in online service delivery will adversely impact business. Of course, you do not have front doors, parking lots, and landscaping with online services, but you do have product catalogs, search and navigation services, and checkout services. How well these function and perform can shape a customer's perception of a business and ultimately how much business she or he does with your company.

This chapter will examine the potential pitfalls of poor service performance along four key business concerns:

- Maintaining and growing revenues
- Acquiring and retaining customers
- Controlling service delivery costs
- Protecting brand image

Let's start by reviewing evidence for the direct link between business metrics and service performance.

Maintaining and Growing Revenues

Time is money. Lagging service performance costs businesses money. Customers search less when performance is poor. Customers abandon shopping carts more when service performance is subpar. No matter what you use for key performance indicators, chances are poor service performance will adversely affect them. This assessment is not just a hunch; it is backed up by a number of studies that measure the impact of Web application performance on user behavior. Let's look at the results of three studies on user experiences from:

- Google
- Bing
- Shopzilla.com

These services performed controlled experiments to measure the impact of service delays to get real data. Anecdotal data cannot answer the question, "How much will slower performance cost me?" These studies can.

Google Slows Search Results

Google conducted a series of experiments on a small percentage of searchers who experienced small, artificial delays in their searching. The results were reported at the O'Reilly Velocity Conference in 2009. They used the following methodology in the experiments:

- A number of users were selected and randomly assigned to a control group or an experiment group.
- The members of the control group did not experience any change in their search services.
- Members of the experiment group experienced delays imposed before headers were displayed, before ads were displayed, and before search results were displayed.
- The length of delays varied from 50 to 400 milliseconds.
- User searches were delayed for varying time periods from 4 to 6 weeks.

At the end of the experiment, Google found that searchers experiencing delays reduced the number of searches they did (that is, they used the service less). Furthermore, the more searching someone did before the delays, the more they reduced their use of the search service. It is not clear whether the users reduced the number of Web searches they did or if they performed some searches using another service.

Bing Slows Things Down, Too

The Microsoft search engine Bing performed a similar experiment with similar findings. As with the Google experiment, a small number of users were selected as subjects of the experiment. Delays of 50 milliseconds to 2 seconds were introduced when returning search results. As one might expect, large delays had significant impact. In fact, the 1 and 2 second delay had such immediate and adverse affects on users that the experimenters stopped the experiment after a short period of time. The sub-1 second experiments continued to run for 3 weeks. Some of the results are shown in Figures 3.1 and 3.2.

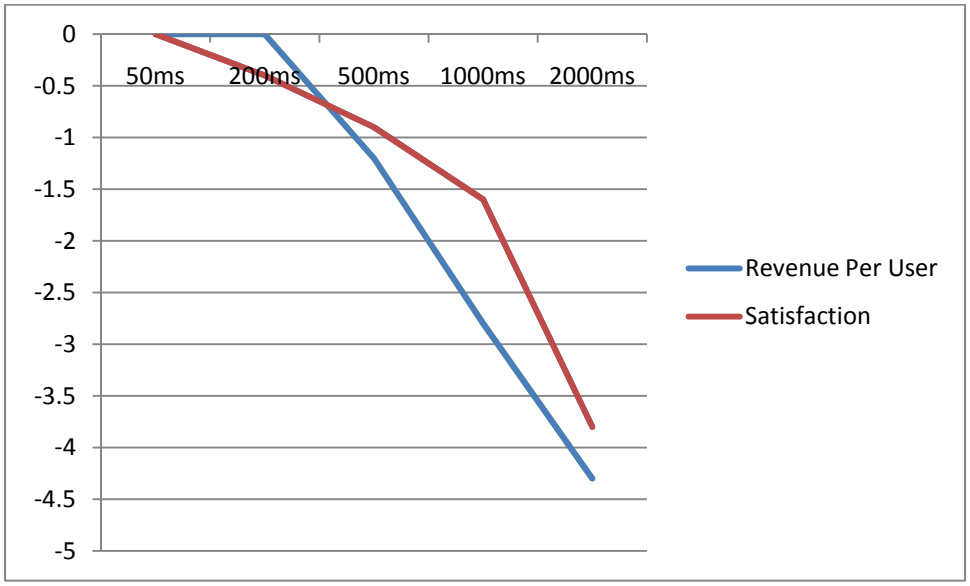

Figure 3.1: Bing study shows accelerated drop in both revenue and satisfaction as performance delay increases. Graph shows percent decline with increase in time to receive results.

Not only do key performance measures, such as revenue per user and satisfaction (a heuristic measure used at Bing) drop, but the rate they drop increases as the performance delay increases.

One particularly interesting measure tracked in the Bing study is the time to click once a page is rendered. Bing found that the increased time to click was roughly twice as long as the delay. Time to click is a proxy for measuring user engagement. This metric shows that the longer the delay, the more disengaged the user becomes.

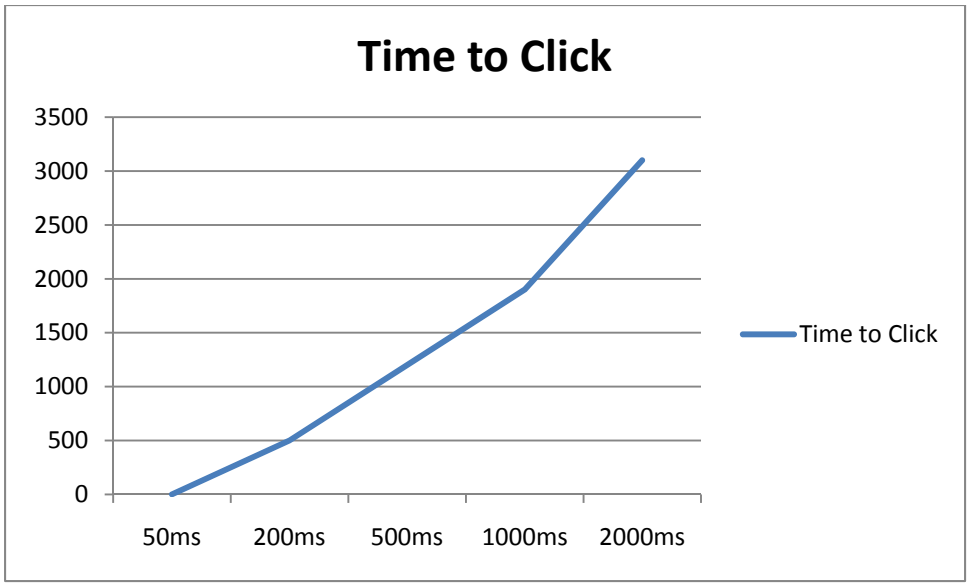

Figure 3.2: Bing study also shows an increase in time to respond to results as the delay in receiving results increases. Graph shows increase in time for user to respond with next action with increased time to receive results. All times in milliseconds.

Another significant finding from these two studies is that the cost of delays increases over time and they persist. The take-away message from these studies is that:

- Performance directly impacts the user's willingness to use and engage with an online service
- Slower performance leads to reduced revenue
- Slower performance leads to less engaged customers
- Once a customer has experienced poor performance, their response to the experience may continue even after performance returns to normal

We could argue that these findings reflect users' expectations for search engines and do not necessarily translate to ecommerce or other online business services. Such is not the case, as another experiment shows.

Resource

For more information about the Bing and Google experiments, see Eric Schuman and Jake Burtlag's presentation "The User and Business Impact of Server Delays, Additional Bytes, and HTTP Chunking in Web Search" at http://en.oreilly.com/velocity2009/public/schedule/detail/8523.

Shopzilla Increases Performance and Revenues

Shopzilla.com is a comparison shopping site with more than 40 million shoppers a month generating 8000 searches per second leading to 100 million impressions a day based on more than 100 million products. The service started as bizrate.com in 2000 and grew from there—and so did page load times. Eventually, page loads started to take 4 to 6 seconds. The company redesigned their core applications and realized the benefit immediately:

- Reduced page load times to 1.5 seconds
- Increased page views by 25%
- Increased conversion rate by 7 to 12%

In the Shopzilla business model, conversion rate is correlated to revenue; a 10% increase in conversion rate translates into a 10% increase in top-line revenue. The key take-away from this experiment: Time is money.

> **Resource**
>
> For more information about the Shopzilla redesign, see Philip Dixon's presentation "Shopzilla's Site Redo—You Get What You Measure" at http://velocityconference.blip.tv/file/2290648/.

The Google, Bing, and Shopzilla.com studies all show a correlation between performance and business metrics. Poor performance hampers and degrades what customers do at a Web site. This leads to reduced use of the site and that in turn results in some cases to quantifiably lower revenues.

The data shows the relationship between performance and business metrics. The next task is to try to understand why this relationship exists.

Acquiring and Retaining Customers

Acquiring and retaining customers is a complex, multifaceted process. Some of the factors that influence customer retention, such as the competitiveness of the market, are outside the control of a single business. Others, such as customer experience and satisfaction with a company's Web site are within a business' control. In this section, let's consider conditions that contribute to customer retention and see the role of external monitoring in giving businesses the data they need to retain those customers.

A Tale of Two Markets

Marketers have the job of building brand recognition and introducing potential customers to a company's offerings. In some markets, acquiring and retaining customers is easier than in others. A couple of examples will demonstrate why online services do not fall into the "easier" category.

Low Competition Markets

In emerging business areas where new services are offered, marketing is like a land rush. There are more opportunities than competitors and often being first is more important than being good. This is especially the case when the first business in a market can exploit market conditions to reduce the chances of competition. For example, consider a rural area without high-speed Internet access. There is probably sizeable pent up demand for high-speed access, so the first company to deliver even a reasonable approximation of high-speed Internet access will do well. Add some short-term incentives, like half price fees for the first 2 months, along with long-term commitments, say a 2-year contract, and the first provider in the area will probably have enough market share to leave competitors with limited ability to recoup their costs if they choose to compete. In markets like this, businesses do not have to attend to customers as much as they would in a competitive market. There is an imposed "stickiness" to this business model.

This example highlights conditions under which customer acquisition and retention need not be a primary business concern. They include:

- Limited options for switching to another product or service provider
- High barriers to entry for competitors, such as the cost of running coaxial or fiber optic cable
- High cost of switching services, such as early contract termination fees or requirement to repay incentives
- Early control of substantial market share by a single product or service provider, which reduces opportunities for competitors

A rational business strategy under conditions such as this is to reduce costs associated with customer retention as much as possible. After all, where are they going to go?

High Competition Markets

Not all markets meet these conditions and the implications for customer acquisition and retention are apparent. Let's consider the marketing and customer management challenges faced by an online retailer.

An online business selling home and garden products faces stiff competition from traditional retailers with physical stores, online-only retailers, and companies with both online and brick-and-mortar outlets. What is it about an online retailer that would make it a more appealing site than any of its competitors? The short list of features likely includes:

- Product availability
- Price
- Convenience
- Quality
- Reliability

There is no point shopping in a store that does not carry what you need, so availability tops the list. Price is important but how important varies by customer and product; commodity products are more price sensitive than specialized services. Many of us would spend more time negotiating a new car price than the fee for heart surgery. When it comes to availability and price, information technology can help but only in a supporting role.

The infrastructure and software behind an online retailer are critical factors determining the convenience, quality, and reliability of customer interactions on the Web site. They are also distinguishing factors that make the online retail market so competitive. Unlike the previous example about Internet access providers, there are few barriers to entry, customers can switch providers with low cost to themselves, and there are competitive options to choose from. In markets such as these, customer retention is more difficult and significant effort is needed to maintain customer satisfaction. One critical element of this process is understanding the user experience on the Web site. For that, we depend on external monitoring. This type of competitive environment is not new to retail but online retail has even lower costs of switching. Customers who feel they are waiting too long in a checkout line may hesitate to leave a cart in a store and drive to a competitor's shop. Online, it is a trivial task to navigate to another site.

How do you know when your site is unresponsive and leading to customer loss? Frequent external monitoring is the only way to assess site responsiveness over a wide geographical area over extended periods of time. We cannot depend on the fact that a site is responsive during a testing phase, or that a site was responsive 2 weeks ago when we ran an ad hoc external test. Constant and comprehensive monitoring can provide early warnings about performance problems before they significantly and adversely affect the customer experience.

Understanding the Customer Experience

The customer's experience at a Web site influences how they use the site, whether they come back, and most importantly, if they are willing to make a transaction while at the site. The overall customer experience can be broken down into three influencing factors:

- Customer expectations
- Customer satisfaction
- Performance and usability

The first two of these are shaped, in significant part, by the third.

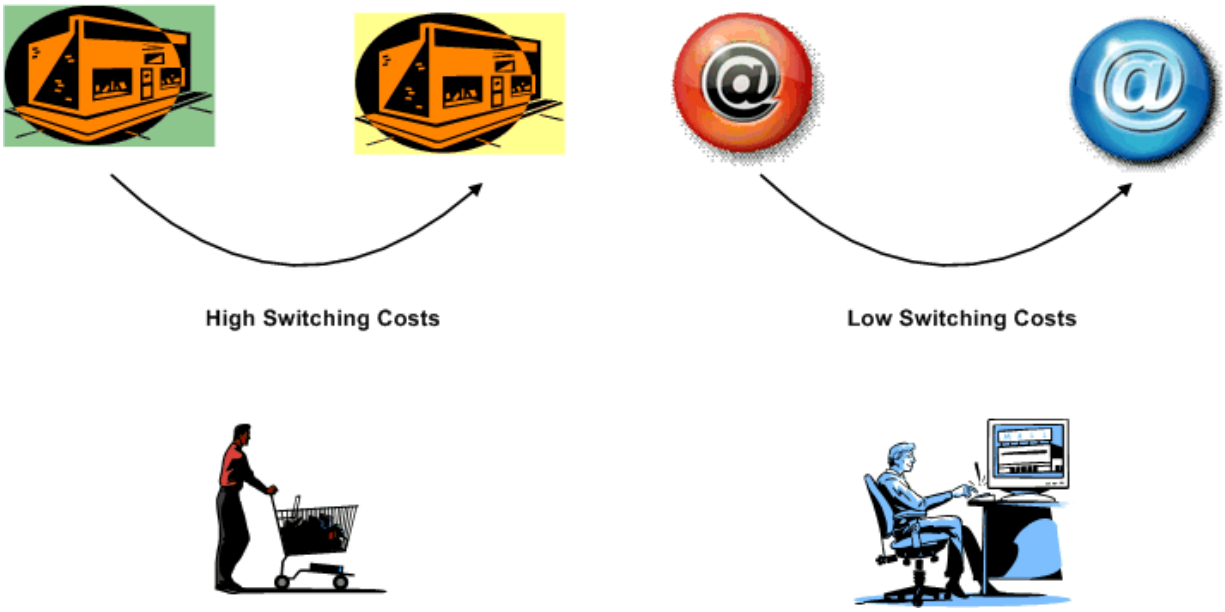

Figure 3.3: The cost of switching providers is lower online than with physical stores making customer retention even more difficult.

Customers bring expectations with them to a Web site. Some of these expectations are shaped by their immediate needs. If I am looking for a textbook, I expect to be able to find it at one of the major online book sellers. If I am using a search engine to research digital camera reviews and an ad is displayed and I click on the ad, I expect to be taken to a relevant page displaying cameras similar to what I was searching for.

In addition to these immediate, short-term expectations, we bring with us more generalized expectations as well. These are longer-term, more persistent expectations that apply to a wide variety of situations. Some examples of Web site–related expectations include:

- Reasonably fast page load times
- Uncluttered pages
- Readily understood navigation schemes
- Minimal number of steps to complete a task

These contextual expectations are shaped, in part, by our experiences with Web sites and services in general. Google's main search page is the hallmark of minimalism in Web page design. There is no clutter whatsoever and the page loads quickly. Retail Web sites and others that have more complex interactions with users require navigation structures to help users get to specific areas within the sites and to know that such areas exist. We see this type of design in online retailers, banks, and other service providers.

Customer satisfaction with these sites largely depends on how well the customer expectations are met. When sites are slow, unresponsive, or difficult to navigate, customer satisfaction will drop. As the previously mentioned Bing study shows, drops in customer satisfaction correlate with drops in revenue.

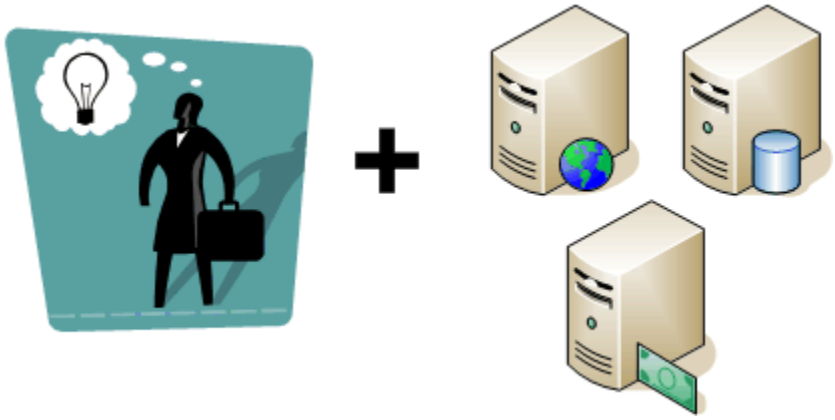

Figure 3.4: Customer satisfaction with a Web site is a function of their expectations and the performance and usability of the Web site.

Customer satisfaction can be increased with improved performance and usability. It is important to remember that performance and usability must be assessed from the perspective of the customer, not an internal testing and quality control group within your organization. This is the reason external Web site monitoring is needed to support customer acquisition and retention.

External Web site monitoring can help identify isolated performance issues. You cannot always make blanket statement like "The average page load time is 2.4 seconds" without qualifying the sentence:

- Where was this measured?
- Over what period of time?
- During what hours of the day and days of the week?
- What parts of the Web site were tested?
- Were page load times weighted according to the frequency with which they are loaded by customers?
- Were the page load times weighted according to the importance of the page to customer abandoning a shopping cart?

You should also consider the fact that averages can be skewed by fast performance on pages that are less important or less frequently loaded than others. Rapid load times on Help pages are desirable but cannot make up for sluggish response on cart display or checkout pages. Details about performance matter, so it is important to use external monitoring to collect data about specific areas that are potential sources of customer dissatisfaction.

Sources of Customer Dissatisfaction

Web site designers, application developers, systems administrators, and network managers all have a role in creating and preserving customer satisfaction. Their jobs often entail addressing common sources of dissatisfaction with Web sites and online services, including:

- Poor performance
- Inaccessible services
- Inconsistent service levels
- Improperly executed transactions

We have seen from the three studies discussed earlier that even marginally poorer performance than normal can adversely impact a business.

Poor Performance

Customers see poor performance in the time it takes pages to load, the time it takes for servers to respond to basic operations, such as searching, and the time required to complete basic tasks, such as checking out. Customers do not necessarily care why a system is slow, but as a service provider, you do need to know. Complex rich Internet applications (RIAs) can be slowed by:

- Inefficient client-side processing
- High latency networks
- Poorly tuned server-side applications
- Improperly optimized databases
- High overhead protocols
- Older Web browsers
- Many images and other large objects on the page

External monitoring services can help identify sources of poor performance. For example, poorly tuned server-side applications will impact customers regardless of their location; high latency networks or other connectivity problems may be more isolated to a subset of customers. Similarly, inefficient client-side processing will adversely impact customers with older or lower-end devices than those with high-end, multi-core, memory-rich workstations. Poor performance is sometimes a global problem with a Web site and sometimes it is isolated to particular pages, backend components, or Web services. A comprehensive set of monitoring scripts can help identify isolated pockets of poorly performing components; we will address this topic in more detail later in the chapter.

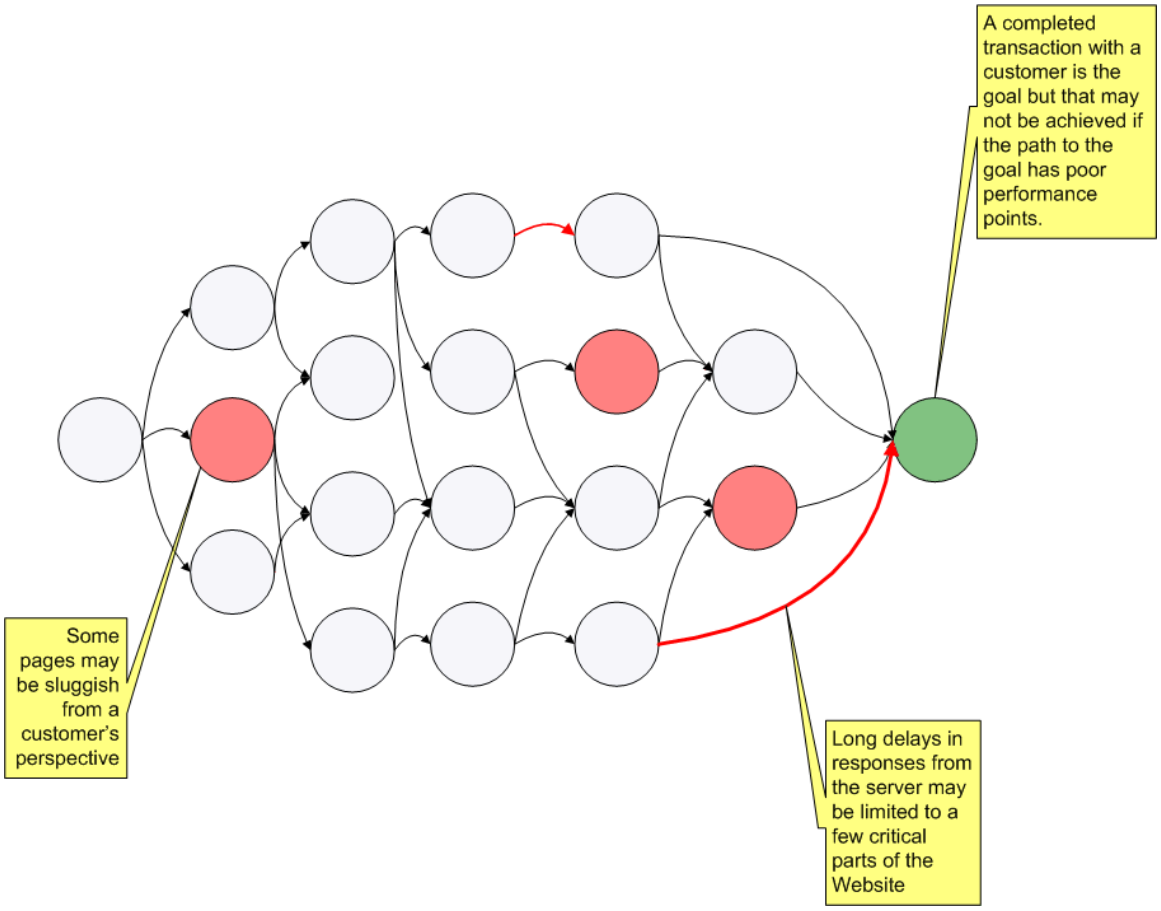

Figure 3.5: Poor performance is not necessarily a global problem. It is important to identify problem areas, especially those that are highly correlated with customers abandoning shopping carts or leaving the site before completing a transaction.

Inaccessible Services and Inconsistent Service Levels

Another contributing factor to customer dissatisfaction is inaccessible services and inconsistent service levels. A critical error, like not being able to access a site, can make a customer call into question the quality of service and the ability of the business to deliver those services. An error message like that depicted in Figure 3.6 might come from a server providing a fairly minor contribution to a transaction, such as displaying a calendar with expected shipping and arrival dates. The service may not even be in your control because it is provided by a business partner. The bottom line is that this inaccessible service is perceived as a problem with your Web site.

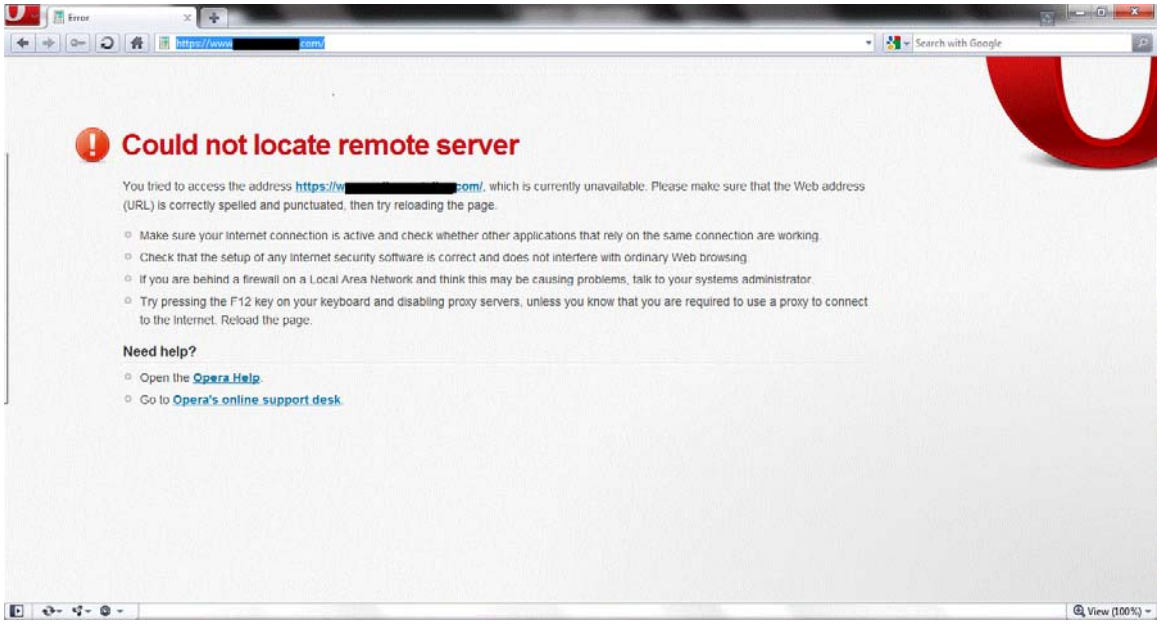

Figure 3.6: With sufficient external monitoring, you can detect problems with a Website before customers are affected.

Inconsistent performance levels can also frustrate customers. Questions such as "Can I make a quick purchase or is this going to be one of those bad days with this site?" can dissuade customers from returning. Again, the cost of switching to another online provider can be so low that even modest performance problems and inconsistencies can prompt customer churn.

External monitoring can help with both inaccessible services and inconsistent performance. In the case of the former, external monitoring can provide early detection of problems. In the case of the latter, external monitoring collects data over extended periods of time that can help identify inconsistent performance.

Improperly Executed Transactions

If there is anything worse than poor performance, it is improperly executed transactions. Imagine spending 30 minutes browsing through a Web site, comparing products, reading reviews, and making selections only to get to the checkout and have an error in the last steps of the purchase. Once again, external monitoring can help identify such problems.

External monitoring scripts can test the full range of Web site functionality. Scripts can methodically test all possible paths through a site and exercise different features on each page. You can even borrow a technique from security professionals known as *fuzzing*. This is the practice of sending unexpected inputs, such as overly long text strings or improperly formatted dates, in an attempt to break a service. Of course, for security professionals, the goal is to identify security vulnerabilities in a system; in this case, you are looking for potential points of failure in customer transactions.

Acquiring and retaining customers in online environments is difficult. Customers are increasingly expecting feature-rich sites and services. To deliver these, businesses are deploying increasingly complex applications. Customers do not have to tolerate sluggish sites or inconsistent service because the cost of switching to another vendor is so low. You cannot change customers' expectations for rich functionality, and you will never design 100% reliable, high-performing applications that can adapt to any possible workload. You can, however, use external monitoring to detect problems early and, as we shall see in the next section, use external monitoring to improve the efficiency of service delivery.

Controlling Service Delivery Costs

External monitoring provides a framework for supporting operations delivery over the long term—it is not just for identifying isolated problems before customers see them. This framework enables reduced labor costs as well as reduced infrastructure costs.

Reducing Labor Costs

External monitoring can reduce the staff resources needed to detect and address performance problems. The benefits come from minimizing troubleshooting and reusing monitoring scripts. In the end, these benefits of external monitoring allow you to solve problems faster and more efficiently.

Minimizing Troubleshooting

The time required to troubleshoot performance problems often depends on the amount of data you have available about a problem. In an environment without an external monitoring program in place, you have to collect data in response to each performance incident as it arises. For example, assume an analyst in sales and marketing notices a drop in transactions over the past week. She browses the Web site looking for clues to the downturn and discovers the site is not performing as normal. She submits a service desk ticket. This is the first indication for IT support that there is a problem.

IT support takes it from here. A technician browses the site and verifies the problem. Of course, browsing alone does not provide enough information to diagnose the problem, so the technician starts to dig into log files—many log files, including Web site, application server, search service, and database logs. This is a time-consuming process. Unfortunately, it is not a comprehensive process; log files do not provide the kind of detailed information about the client that is needed to diagnose performance issues with RIAs.

Now consider the same situation in a business that uses external monitoring. Monitoring scripts are run on a regular basis from multiple external test nodes, so there exists a comprehensive set of data about performance. Rather than taking a week for someone to notice the effects of poor performance (the drop in transactions), the external monitoring reports would have highlighted specific problem areas. IT staff could use the data to better understand the nature of the performance problems by considering questions, such as:

- Are all pages affected or only some?
- Do the problems appear to all customers or are they isolated to a specific geographic region?
- What types of actions, such as database queries, services calls, and so on are experiencing the greatest delays?

The external monitoring data can help IT professionals narrow the range of possible problems and hone in on a solution without having to run ad hoc tests of their own.

Reusing Monitoring Scripts

A comprehensive monitoring regime could become a time-consuming and high-cost overhead expense if it were not for the way you develop and deploy monitoring scripts. These scripts are highly reusable because a template can provide core functionality, and with more specialized instances, derived from templates. For example, a test engineer might define a generic purchase script that:

1. Searches for a product
2. Adds a product to a shopping cart
3. Repeats steps 1 and 2 a predetermined number of times
4. Views shopping cart
5. Selects from alternate payment and delivery methods
6. Concludes the order

The template can be customized to test particular types of products, used for testing boundary conditions (for example, adding 100 items to a cart), and testing all possible options, such as testing each payment method. Each of these specialized scripts can then be deployed to multiple test nodes in a variety of geographic areas. Comprehensive testing need not be time consuming or costly when scripts are reused.

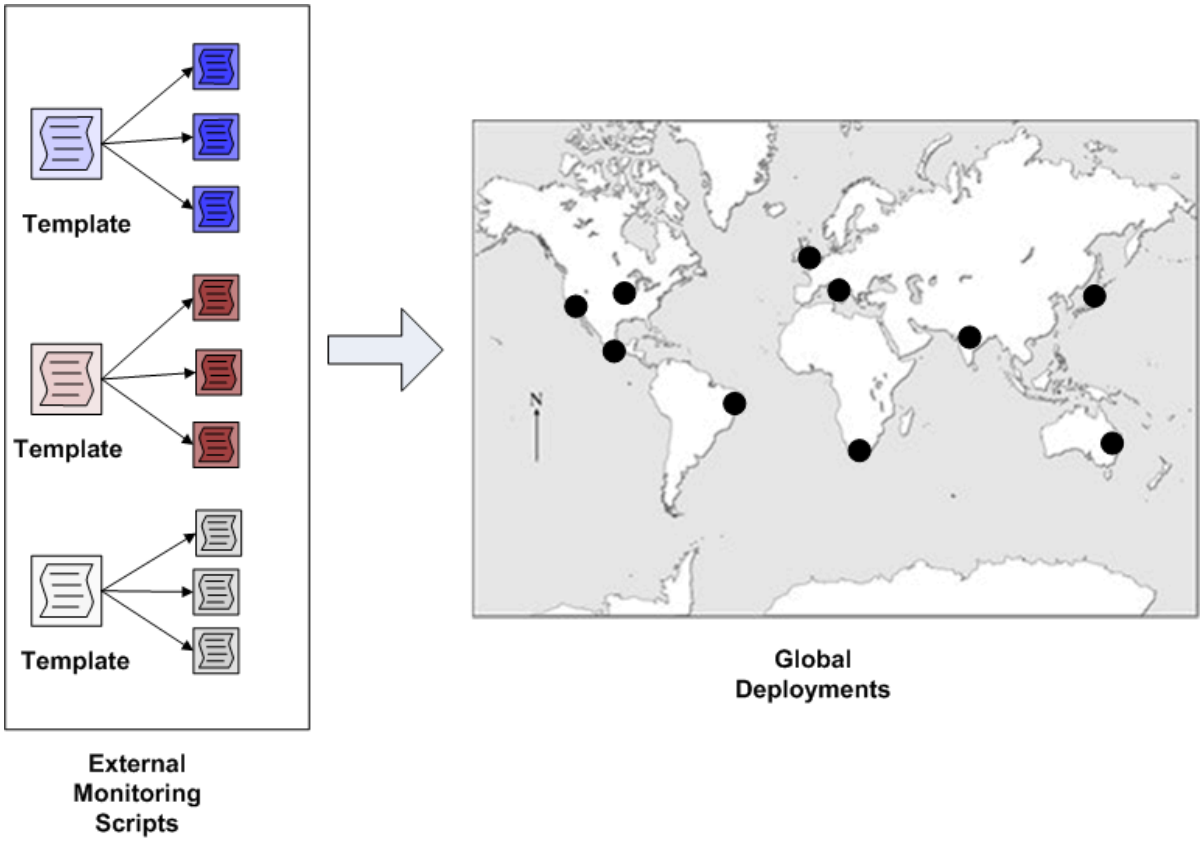

Figure 3.7: A small number of templates can be reused to define more specific monitoring scripts that, in turn, can be deployed to multiple external locations around the world.

Reducing Infrastructure Costs

The data collected by external monitoring can also help better manage infrastructure. The data describes how customers experience Web-based operations. This data can be even more valuable when coupled with internal performance monitoring data.

For example, external monitoring data may show a slowdown in performance for some transactions at the same time database monitoring software indicates high memory utilization in the database server. Together, these two pieces of data justify adding more memory to the server to address the performance issue. Alternatively, there may be a stored procedure within the database that is inefficient and needs redesign. Database administrators could review data on which procedures or SQL statements were running at the time of the performance lag.

A constant concern for systems administrators is providing high availability. If a server goes down, other servers in a cluster or a backup server should take over the workload automatically. Architects have a number of options for providing failover protection. They could couple every critical server with a dedicated backup server, they could put a number of servers in a load-balanced cluster, or they could implement a failure detection/rerouting scheme that forwards traffic to another server. The best option will depend on a number of considerations, including:

- The load on servers and performance levels
- The geographic placement of servers
- The network bandwidth and latency between servers
- Impact on customers and revenue if a server is down

Once again, external monitoring data can provide valuable insight into the detailed functioning of applications. If external monitoring data indicates network capacity problems, failing over to a server in a different geographic area would not be a good option; running multiple servers in a load-balanced configuration may be better. If the monitoring data indicates server performance is near capacity, running a load-balanced cluster could help with both performance and high availability.

A Note on In-House vs. Service-Based External Monitoring

It should be noted that the kind of detailed information provided by external monitoring is not available from internal monitoring. Of course, internal monitoring is valuable; it can provide details that external monitoring cannot. For example, detailed database performance monitoring about CPU and memory utilization along with statistics on I/O operations is gathered with internal processes and is invaluable for tuning database performance. A comprehensive monitoring program should include both internal and external monitoring. These two types of monitoring complement one another; they do not replace each other.

External monitoring may be conducted in-house or by using a third-party service. The advantages of an in-house approach is that you are in complete control over the location of monitoring servers, the types of monitoring jobs that are run, the schedule at which they are run, and so on. The disadvantages are fairly clear, too.

You have to deploy and maintain remote servers. To ensure adequate coverage, you need to deploy external monitoring servers to multiple locations. If you are working with a global operation, then monitoring should have a global reach as well. Also, there is the overhead of maintaining a monitoring site. You have to be careful to not run your monitoring service from a site that is not representative of your customers in the area. For example, you should not execute monitoring scripts from a data center with higher bandwidth networking than is available to your customers; the results will not be representative of a customer's experience.

External monitoring services should have the building, networking, and server infrastructure broadly distributed to provide sufficient geographic coverage. They should also provide support for script development, deployment, and execution along with reporting services. As with any specialized service, they should be able to deliver these services at a lower cost because of specialization, decreasing marginal costs of providing the services and related factors. A potential disadvantage is decreased control over implementation issues. Concerns about implementation often have to do with the need to ensure adequate performance, security, reliability, and so on, although these concerns can often be addressed with service level agreements (SLAs).

Protecting Brand Image

One of the major drivers for adopting external monitoring is to protect brand image. The value of a brand has become more understood over the past several decades through concepts such as brand equity (the value of a brand) and brand extension (the ability to leverage the value of a brand to introduce new unrelated products). A brand is understood today to be an asset even though it is not something we store like inventory or deploy like infrastructure. As a valued asset, brand is something to be protected against threats. One threat facing brands is the association with poor performance.

Consider some of the ways poor online application performance can adversely impact brand:

- Customers may associate poor application performance with poor product quality
- Poor performance may be coupled with negative views about a company's ability to deliver products and services
- Poor performance online may be generalized to other channels, such as physical retail centers
- Reseller and business partners may have their brands damaged by association with your poorly performing site

The best way to avoid the potential brand damage associated with poor application performance is to maintain consistent service levels. This starts by deploying adequate infrastructure, designing applications that perform well on typical customer hardware, and monitoring performance both internally and externally.

Many of the benefits outlined earlier with regards to acquiring customers and controlling costs apply to brand protection as well. Rapidly detecting and responding to performance problems reduces the overall impact of those problems. Poor application performance can contribute to customer churn, increase maintenance costs, and adversely affect brand image.

The cost of acquiring and retaining customers and the costs associated with additional maintenance may be felt relatively quickly. Brand damage, however, can occur more slowly and the effects may linger for even longer periods of time. Just as it takes a great deal of time and effort to build a brand, it takes significant resources to rebuild it after it has been damaged. Fortunately, negative associations with a brand because of poor application performance are some of the most readily controlled threats to brand value.

Summary

External monitoring is not an end in itself. Its value stems from the contributions it makes to acquiring and retaining customers, controlling operational costs, and protecting brand value. Studies by leading online firms provide clear evidence that performance is linked to key business performance metrics, including revenue. Online businesses are subject to unusually low switching costs making it easier for customers to move to competitors. For businesses, this means they have less margin for error when it comes to any factor that might drive a customer away. Poor application performance is one of those factors.

External monitoring plays a crucial role in providing early warnings about performance problems. These systems also collect valuable data that complements internal monitoring data. As a result, businesses can more effectively track their applications' performance, respond to problems, and mitigate the risk of unwanted customer loss, increased maintenance costs, and brand damage.